Forensic Science R&D Program
National Institute of Justice
Department of Justice

Award No.:

2014-DN-BX-K007

Project Title:

Adult Age Estimated from New Skeletal Traits and Enhanced Computer-Based Transition Analysis

PI: George R. Milner
Professor
Department of Anthropology
The Pennsylvania State University
University Park, PA 16802
ost@psu.edu
814-865-2509

Co-PI: Stephen D. Ousley
Adjunct Professor
Department of Anthropology
University of Tennessee
502 Strong Hall
1621 Cumberland Avenue
Knoxville, TN 37996-1525
craniator@statsmachine.net

Co-PI: Jesper L. Boldsen
Lektor
ADBOU
University of Southern Denmark
Campusvej 55
DK5230 Odense M Denmark
jboldsen@health.sdu.dk
+45 6550 2603

John Hanold, Associate Vice President for Research; Director, Office of Sponsored Programs
osp@psu.edu -- Phone: (814) 865-1372

Submission Date: July 19h, 2018

DUNS EIN Numbers

The Pennsylvania State University
Office of Sponsored Programs
110 Technology Center
University Park, PA 16802

Grantee IRS/Vendor No.

Grantee IRS/Vendor No.

Project Period From 01-01-2015 to 12-31-2019

Reporting Period End Date: 12-31-2019

Final Summary Report? Yes

Introduction

For decades it has been widely recognized that methods for aging adult human skeletons yield estimates that are inaccurate, imprecise, and biased; it is also impossible to estimate the ages of old people (ca. 50+ years). Forensic anthropologists consequently have a difficult choice to make: either use methods known to yield poor results or simply rely on their experience, with all that implies in terms of reliability, replicability, and courtroom admissibility.

This NIJ project directly addresses the long-standing problem of adult skeletal age estimation by adopting a radically different approach from all other studies. It features three components, starting with new skeletal age indicators, moving to a refinement of analytical procedures, and ending with a user-friendly computer program. Previously published studies focus on one or, perhaps, two of these three pressing needs. No previous research effort has attempted to tackle all three objectives simultaneously and in an integrated fashion.

In addition to the development of a workable procedure for estimating age throughout adulthood, the project has a larger objective. It is to reorient the entire process of skeletal age estimation, starting with the kinds of traits used and extending through the ways they are combined to yield ages tailored to the characteristics of individual skeletons.

Skeletal Traits

The project is based on a new set of skeletal traits, many never examined systematically before. Most of the traits were defined after years of preliminary work with multiple diverse skeletal collections that demonstrated the value of relying on experience when estimating age. An experience-based approach, however, has flaws that render it unacceptable for forensic investigations. It cannot be taught except through a lengthy apprenticeship period, and there is no means to estimate quantitatively an observer's (un)certainty about age estimates. This project,

therefore, disaggregated what is seen by the expert eye, defined those traits, and systematically assessed them for their effectiveness as age indicators.

The intent from the beginning was to base the NIJ method on many binary traits (e.g., absent or present). Doing so facilitates data collection, reduces observer error, and simplifies learning the procedure. Traits were identified that undergo transitions from early to late stages at different points during life, so age can be estimated throughout adulthood.

Data were collected from well-documented known-age skeletal collections to obtain a large and diverse sample: the Bass Donated Collection at the University of Tennessee, donated skeletons at Mercyhurst University in Pennsylvania; the Pretoria Bone Collection at the University of Pretoria in South Africa, the Chiang Mai Collection at Chiang Mai University in Thailand, and the Lisbon Collection at the Museu Nacional de História Natural e da Ciência in Portugal (Table 1). To date, 1,774 skeletons have been examined, 1027 from males and 747 from females (Table 1).

Table 1. Collections, Samples, Skeletal Traits, and Age-at-Death Transitions*

		Skeletons			**Traits**			
Collection	**Date**	**N**	**M**	**F**	**New**	**Used**	**Elim**	**Trans**
Bass	05/15	423	221	202	---	89	15	129
Pretoria	07/15	424	269	155	3	77	1	107
Chiang Mai	01/16	440	271	169	3	79	0	118
Bass (return)	02/16	14	12	2	0	79	0	118
Mercyhurst	04/16	7	5	2	0	79	0	118
Lisbon	06/16	375	183	192	0	79	0	118
Pretoria (return)	08/17	91	66	25	0	79	0	118

N	Number
M	Male
F	Female
Elim	Eliminated
Trans	Transitions

* Final project totals for individuals and traits will change through the elimination of skeletons with questionable documentation, the omission of uninformative traits, and the combination of stages to facilitate accurate scoring.

Skeletons from four continents increase the sample's diversity. That enhances the method's forensic applicability, especially in the United States where people trace their ancestry back to many parts of the world. A large sample is also necessary to pick out an age-related signal from the noise of human variation.

For most forensic applications, there is a need for a global, or universal, age estimation procedure. That is because the original reference samples and the skeletons investigators examine will always differ in their genetic backgrounds and activity levels. The aim was to develop a procedure with wide applicability, regardless of the forensic context.

Table 1 also summarizes changes in the number of traits that took place during data collection. Analyses early in the project allowed the research team to focus on skeletal traits that yield the most information about age at death. Doing so facilitated efficient and cost-effective data collection.

Analytical Procedures

For each age indicator, the age of transition from one stage to the next was established following early work by two project participants (Boldsen et al., 2002). Most of the stages are defined as binary to speed data collection and to reduce observer error by eliminating minor and hard to distinguish differences in morphological details.

Two routes were pursued for combining information to yield age estimates individually tailored to specific skeletons. One is an elaboration of existing Transition Analysis (TA^2)

(Boldsen et al., 2002). The other involves machine learning, which has already yielded promising results.

Computer Programming

The means of producing accurate age estimates with numeric and graphical output for individual skeletons requires computer programming. It is essential for the success of the research endeavor because no matter how good a method might be, it has little practical value if it cannot be readily used by forensic practitioners in both laboratory and field settings. A computer program has been produced to expedite data entry and analysis, which will be available this spring, and it will continue to be developed based on feedback from users and advances in analytical methods. Software for data entry and analysis, as well as extensive documentation, will be made freely available to forensic practitioners.

Results from Conventional Methods

As a baseline to measure this project's performance, over 200 American males and females from the Bass Collection were aged using the two most frequently used age-estimation methods employed by forensic anthropologists (Garvin and Passalacqua, 2012). One focuses on the public symphysis (Brooks and Suchey, 1990), and the other on the iliac portion of the sacroiliac joint (auricular surface) (Lovejoy et al., 1985). Figure 1 compares the actual ages of skeletons to estimates from the original publications (dots): means for the Suchey-Brooks pubic symphysis stages, and range midpoints for the Lovejoy sacroiliac stages. In the absence of information about the age composition of the original reference sample, the age used for the Lovejoy terminal, open-ended interval (60+ years) is the median age of individuals in our sample. Also shown is the identity line (solid) and the loess trend of estimates (dashed).

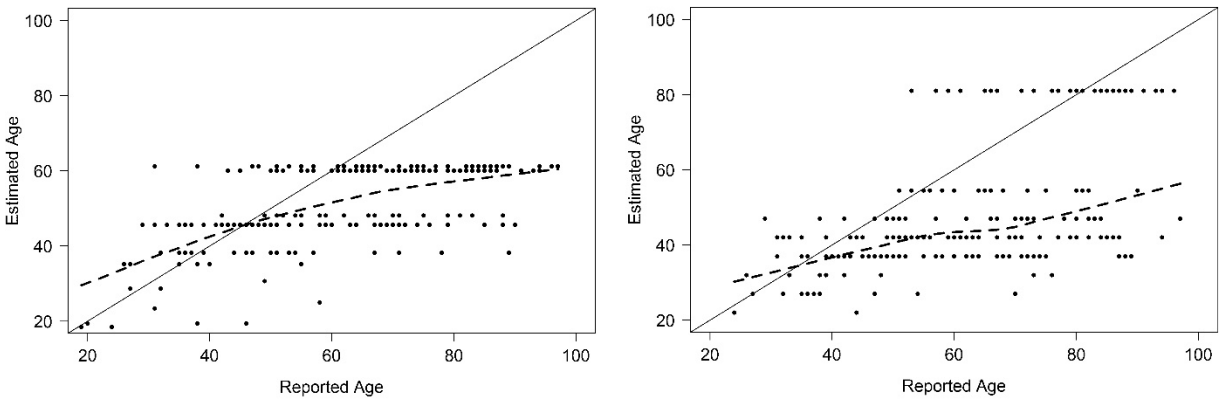

Figure 1. Reported and estimated ages (dots) for a sample of American skeletons, with the latter from the Suchey-Brooks pubic symphysis (left, N=222) and the Lovejoy iliac auricular surface (right, N=213) methods. Also shown are the identity line (solid) and a locally weighted loess curve (dashed).

If age was estimated without error, all individuals would line up on the diagonal. For both standard methods, there is not only a poor correspondence between estimated and true ages, from middle adulthood onward ages are underestimated.

Results from the New NIJ Method

After decades of work – including new ways to classify age-related skeletal changes (e.g., Boldsen et al., 2002) – it is reasonable to believe that the anatomical structures that have received the most attention, notably the pubic symphysis and sacroiliac joint, have yielded about all that they can for age estimation. In contrast, experience-based age estimates can be remarkably accurate when undertaken by researchers who have examined many hundreds of known-age skeletons that collectively span all of adulthood (Milner and Boldsen, 2012: Fig. 14). Because such estimates are not obtained through a fixed procedure with known error that can be readily taught to others, they cannot be a central part of forensic investigations where results must be replicable and have an objective basis with quantifiable certainty. The NIJ project was designed to disaggregate and quantify what experts see when they examine skeletons.

For each bony trait, a transition curve is calculated where the probability at each age of undergoing a change from one stage to the next is estimated. The simplest transition curves are S-shaped so that the more advanced (older) skeletal feature simply increases monotonically throughout adulthood. In other words, the proportion of people in the mortality sample with the skeletal characteristic increases with advancing age. Three such transition curves are shown in Figure 2. The most informative for age estimation are steep curves, which are concentrated in early adulthood. Most traits have curves similar to the slope of the one in the middle plot. Other traits occur only infrequently in our sample, even in old age.

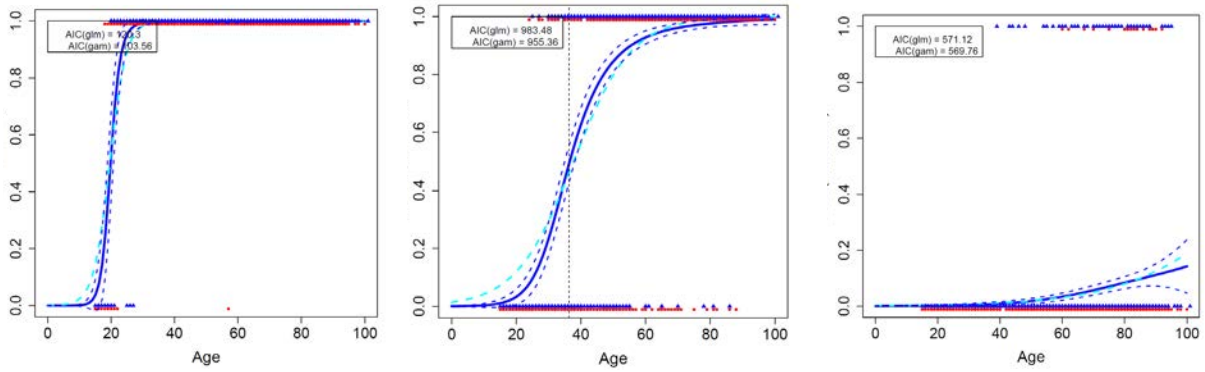

Figure 2. Transitions from one stage to the next vary from rapid (left) to those where a trait has a low probability of occurring, even in old age (right). The x-axis is age and the y-axis is the probability of a trait being present.

Traits that are not useful include those where reasonably flat probabilities of having the skeletal feature span much of adulthood. These bony features are either uninformative about age, or incorrect scoring thresholds were employed that mask any age-related information they might otherwise have provided. Other age-at-death curves are sinuous (locally negative slopes producing bumpy curves); they are, in fact, potentially indicative of selective mortality. Those particular traits cannot be used for age estimation because peaks somewhere in adulthood, followed by a decline, equal the probability of having the old form of the trait much later in life.

This resource was prepared by the author(s) using Federal funds provided by the U.S. Department of Justice. Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Two approaches to age estimation, designated TA^3, are currently being explored. The first is an elaboration of existing TA^2, which involves estimating age from combining transition curves for multiple parts of two pelvic joints and cranial sutures (Boldsen et al., 2002). In TA^3, ages can be estimated through a combination of transition curves, notably those similar to the ones shown in the first and second plots in Figure 2.

The other approach uses machine learning (ML) procedures, such as Random Forests, to select the most informative traits, and Random Generalized Linear Models to predict age using those traits (Hastie et al., 2009; Song et al., 2013). The most promising methods investigated so far are refinements and extensions to linear regression, including the Generalized Linear Model, Principal Component Regression, Partial Least Squares Regression, Quantile Regression, Regression Trees, and Bayesian Regression Modeling. Nearly every method investigated to date can predict age far better than any other published method.

Age estimates from one such ML method (the Lasso) are shown in Figure 3 where the most informative 20 traits are identified and used to estimate age. Although only a simple, semi-automatic procedure, it outperforms traditional methods, including all those used in medicolegal investigations today, as is shown by comparing Figure 3 to Figure 1. Notably the age estimates do not get progressively worse with advancing age, resulting in an ever-widening distribution, and point estimates approximate the identity line.

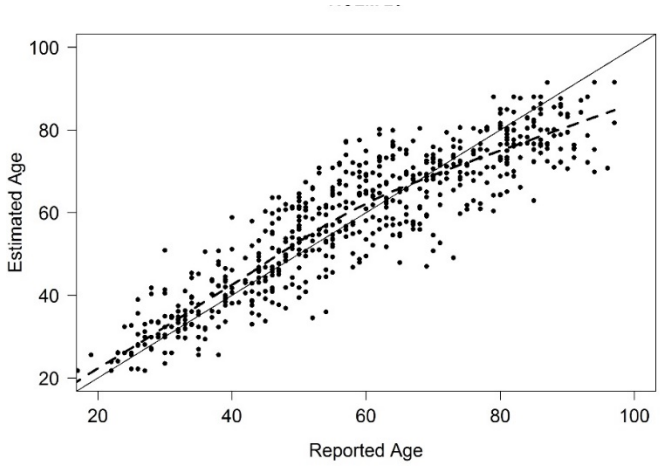

Figure 3. Age estimates (dots) are shown for 606 male and female skeletons with complete observations drawn from a worldwide sample using the most informative 20 traits, as are the identity line (solid) and a locally weighted loess curve (dashed).

Dissemination of Results

Information about the rationale, progress, and results of this project has been distributed through publications, including a published journal article, a submitted book chapter, and two two doctoral dissertations. Numerous lectures and workshops have been delivered to domestic and international audiences. They include two workshops at the 2019 annual meetings of the American Academy of Forensic Sciences (AAFS) and American Association of Physical Anthropologists (AAPA), which rapidly filled to capacity. Feedback indicates the skeletal traits are easy to learn, which reduces observer error, and comments have helped refine the trait-scoring manual, a draft of which was distributed to the AAFS and AAPA session participants.

Project outcomes also include a computer program, a manual covering the scored skeletal characteristics, and a data-collection form. The computer program is available as a beta version, and it will be enhanced through further analyses by the research team and user feedback. The bone-scoring manual, currently 82 pages long, features succinct trait definitions, diagrams, and

illustrations, and it is accompanied by a data-collection form to facilitate the accurate recording of observations by forensic practitioners.

Outlook for the Future

Perhaps the most important outcome of the project is an anticipated shift from observing a few parts of the skeleton for age estimation to a much wider array of bony features. Methods have traditionally focused on skeletal traits that are information rich in the sense that they pass through several age-related changes during adulthood. But as age indicators they perform poorly (Fig. 1). This project has shown that it is necessary to switch to a suite of skeletal traits that individually contribute little to estimating age – they are simply young or old for a particular feature. Collectively, however, they provide much better age estimates (Fig. 3).

Not only are superior age estimates produced, it is possible to extract information about age from incomplete skeletons. That is important because skeletons from forensic settings are often incomplete, so investigators must make do with whatever happens to be present. With traditional methods, age estimates are not possible when key elements, notably the pelvis, are missing. In the new TA^3 approach, age-informative bony features are distributed throughout the skeleton. It is thus possible to estimate age from skeletons missing some, or many, bones.

Conclusion

The immediate goal of the NIJ project was to improve adult skeletal age estimation through the development of a new procedure. A larger objective, derived from the first, was to reorient the field to a new suite of bony traits and a means of quantitatively tailoring age estimates to individual skeletons by combining data from multiple bones. The first objective was met (Fig. 3). The enthusiasm of AAFS and AAPA workshop participants is a sure sign that forensic practitioners are receptive to the second objective. In short, better methods coupled with

better skeletal traits and better initial samples – all of which were part of the NIJ project – inevitably yield better results.

References

Boldsen JL, Milner GR, Konigsberg LW, Wood JW. (2002). Transition analysis: A new method for estimating age from skeletons. In RD Hoppa, JW Vaupel (Eds.), *Paleodemography: Age distributions from skeletal samples.* pp. 73-106. Cambridge: Cambridge University Press.

Brooks ST, Suchey JM. (1990). Skeletal age determination based on the *os pubis*: A comparison of the Acsádi-Nemeskéri and Suchey-Brooks methods. *Hum Evol,* 5, 227-238.

Garvin HM, Passalacqua NV. (2012). Current practices by forensic anthropologists in adult skeletal age estimation. *J Forensic Sci,* 57, 427-433.

Hastie T, Tibshirani R, Friedman JH. (2009). *The elements of statistical learning*. 2nd ed. New York: Springer.

Lovejoy CO, Meindl RS, Pryzbeck TR, Mensforth RP. (1985). Chronological metamorphosis of the auricular surface of the ilium: A new method for the determination of adult skeletal age at death. *Am J Phys Anthropol,* 68, 15-28.

Milner GR, Boldsen JL. (2012). Transition analysis: A validation study with known-age modern American skeletons. *Am J Phys Anthropol,* 148, 98-110.

Song L, Langfelder P, Horvath S. (2013). Random generalized linear model: A highly accurate and interpretable ensemble predictor. *BMC Bioinformatics*, 14, 5.

DISSEMINATION OF RESULTS

Presentations

The largest, and most important, component of dissemination pertained to the transmission of information about the project to colleagues. They included students and professionals, many of which were forensic researchers and practitioners.

During the data collection phase, workshops were presented at all institutions that hosted the NIJ research team. In these presentations, audiences consisted of about 30 or more people. The rationale behind the ongoing NIJ-funded work was discussed in these events, as was the likely outcome of the project. Even though the emphasis in these workshops was on the existing TA^2 procedure (Boldsen et al., 2002; Milner and Boldsen, 2012), such presentations familiarized audiences with how and why transition curves for skeletal traits are used to generate individualized age estimates for skeletons. Thus, the eventual acceptance and knowledgeable use of the NIJ-funded procedure for medicolegal purposes were enhanced through these presentations.

During 2019, the final year of the project period, considerable effort was spent disseminating information about ongoing work to professionals and advanced graduate students, mostly the former. Formal day-long workshops were presented at the annual meetings of the American Academy of Forensic Sciences (AAFS) and the American Association of Physical Anthropologists (AAPA). The workshops included formal lectures and hands-on components with known-age skeletal remains. Both were filled to capacity with 30 participants apiece (the workshops were quickly filled and had lengthy waiting lists). Most attendees were from the United States, although individuals from foreign institutions were also represented, doing much to ensure an international reach to the work.

Especially during the last year of the grant period, project participants presented multiple lectures and workshops at no cost to the NIJ project. The personnel responsible for presentations in the United States, Denmark, and Sweden included Drs. Milner, Boldsen, Ousley, Weise, and Getz. Presentations mostly took place at universities, but also the Defense POW-MIA Accounting Agency (DPAA) in Omaha, Nebraska. Furthermore, Dr. Getz incorporated NIJ results in her part of the Core Principles and Practices in Forensic Anthropology course at Mercyhurst University. Twenty-four people, professionals and graduate students, attended the Mercyhurst course; all were from the United States except for one from Australia. Drs. Milner, Boldsen, and Getz have incorporated NIJ results into teaching in formal (classroom) and informal (hands-on laboratory) contexts for undergraduate and graduate students at the Pennsylvania State University, Idaho State University, and the University of Southern Denmark.

Formal professional meeting podium presentations and posters were presented at the AAFS and AAPA annual meetings (Milner et al, 2016a; Milner et al, 2016b; Ousley et al., 2017; Ousley et al., 2019). They included a presentation at the NIJ R&D Symposium held annually at the AAFS conference.

Through these efforts – workshops, lectures, podium presentations, and posters – the NIJ project results are already gaining attention with a domestic and international audience. Feedback from these presentations, notably the day-long AAFS and AAPA workshops early in 2019, has directly benefited the project, especially by identifying aspects of the data-recording manual that require clarification.

Taken together, these presentations have introduced colleagues, both professionals and students alike, to the rationale of what was done, especially how and why age-estimation in a forensic context benefit from the new approach pioneered by the NIJ project. Our intent from the beginning was to make the development of this age-estimation procedure as transparent as possible for the forensic community, and thereby set the stage for future work by the widest array of researchers possible.

Computer Program, Trait Manual, Trait Scoring Sheet

Publication of the project materials include the computer program, a manual covering the scored skeletal characteristics, and a bone-scoring sheet.

The computer program is available as a beta version, and it will be enhanced through additional analyses by the research team and feedback from users.

The bone-scoring manual, currently 82 pages long with succinct trait definitions, diagrams, and illustrations, deserves special attention because it is essential for the proper use of the procedure by forensic practitioners. Considerable effort was devoted during the project period to getting the manual right. It evolved through a lengthy process prior to, during, and after the first data-collection trip by the research team. The current manual is deceptively simple insofar as definitions are boiled down to their critical components. Excessive, vague, and imprecise wording have been ruthlessly eliminated. Multiple illustrations capture a sense of the range of skeletal variation that investigators might encounter, not just cases that best exemplify a particular trait. That is, an effort was made to show what is likely to be seen when scoring skeletal traits, hence reduce ambiguity in the trait-scoring process.

The manual was distributed to the AAFS and AAPA workshop participants, and it has benefited from their input (Appendix 2 shows one page from the manual). Once again, the project was designed to be as inclusive as possible by involving researchers who work with either forensic or archaeological skeletal remains.

A two-page recording form, designed for rapid data collection, accompanies the manual (Appendix 1). The form has been refined throughout the project period, with the layout evolving to facilitate rapid data collection with a minimum of recording error. The layout of the recording form was designed to speed up both the initial data collection and later computer data entry, as well as to reduce the chance of erroneous trait scores.

Publications

Two research team members – Dr. Getz and Dr. Tarp – completed their doctoral research on related topics, hence benefited directly from participation on the NIJ project. Dr. Getz used the NIJ-generated scoring procedure as the basis for her doctoral research at additional European collections in Greece and in the United Kingdom (Getz, 2017). Dr. Tarp folded data on the NIJ femoral traits into his doctoral work (Tarp, 2017) . Dr. Getz has begun an academic career as a forensic anthropologist in the United States. Dr. Tarp, a Danish national, has secured a post in law enforcement in Denmark.

To date, one journal article based on data collected during this project has been published (Milner et al., 2018b) and a book chapter has been submitted, in addition to the two doctoral dissertations mentioned above. The published journal article focused on DISH, drawing attention to the increased mortality risk experienced by middle-aged American women with the condition. Focusing on selective mortality, the article is an outgrowth of the trait-screening process. As explained above, selective mortality complicated analyses, and it likely contributes to the poor performance of traditional means of estimating the age of skeletons from forensic contexts. The book chapter has been submitted to a volume on age estimation in forensic contexts, and it has an expected publication date as early as December 2020). It covers the origin, rationale, implementation, and initial results of the NIJ project.

Journal articles are planned that address various aspects of the work, and a book that serves as an overview of the entire research process is anticipated. All project participants will contribute to these various publication efforts.

PUBLICATIONS, CONFERENCE PAPERS, AND PRESENTATIONS

Various forms of communication with the forensic community during the grant period – lectures, workshops, publications, and computer program with supporting documentation – are described in the previous Dissemination of Results section. Publications (one published and the other submitted) and professional conference presentations and workshops (i.e., those listed in professional meeting abstracts) are provided below.

Publications

n.d. Milner GR, Boldsen JL, Ousley SD, Getz SM, Weise S, Tarp P. Great Expectations: The Rise, Fall, and Resurrection of Adult Skeletal Age Estimation. In *Remodeling Forensic Skeletal Age*, Algee-Hewitt BFB, Kim J (eds.). Elsevier, Amsterdam, submitted.

2018 Milner GR, Boldsen JL, Ousley SD, Getz SM, Weise S, Tarp P, Steadman DW. Selective Mortality in Middle-Aged American Women with Diffuse Idiopathic Skeletal Hyperostosis (DISH). *PLOS One* 13:e0202283. [doi.org/10.1371/journal.pone.0202283].

Conference Workshops (entire day presentations)

2019 Presenters Boldsen JL, Ousley SD, Milner GR, Getz SM, Weise S, Gampe J; Organizers Milner GR, Ousley SD. An Integrated Approach to Adult Skeletal Age Estimation and

Paleodemographic Reconstruction: Going From Bones to Individual Ages and Mortality Patterns. American Association of Physical Anthropologists, Cleveland, OH.

2019 Presenters Ousley SD, Milner GR, Boldsen JL, Getz SM, Weise S; Chairs Ousley SD, Milner GR. Adult Skeletal Age Estimation: Transition Analysis Using the Entire Skeleton. American Academy of Forensic Sciences, Baltimore, MD.

Conference Presentations

2019 Ousley SD, Milner GR, Boldsen JL, Getz SM, Weise S. A New Method for Adult Skeletal Age Estimation Using Transition Analysis: TA3. American Academy of Forensic Sciences, Baltimore, MD.

2018 Weise S, Boldsen JL, Getz SM, Milner GR, Ousley SD, Tarp P. A Revised Transition Analysis Method to Estimate Age-at-Death from Human Skeletons. European Anthropological Association, Odense, Denmark.

2018 Milner GR, Boldsen JL, Ousley SD, Getz SM, Weise S, Tarp P. Adult Skeletal Estimation: Dismal Results, Ensuing Misconceptions, but a Promising Future. New Methods in Skeletal Age Estimation for Diverse Populations. Stanford University, Stanford, CA.

2017 Boldsen JL, Milner GR. Skeletal Ageing and Selective Mortality. MaxNetAging Conference (Longevity: Future, Present, Past and Across the Tree of Life). Max Planck Institute for Demographic Research, Rostock, Germany.

2017 Milner GR, Boldsen JL, Ousley SD, Getz SM, Weise S, Tarp P. Adult Skeletal Age Estimation: Tackling Long-Standing Problems with a New Approach. American Academy of Forensic Sciences, New Orleans, LA.

2016 Milner GR, Boldsen JL, Ousley SD, Getz SM, Weise S, Tarp P. Estimating Age from Adult Skeletons: New Directions in Transition Analysis Using a Wide Array of Traits. American Association of Physical Anthropologists, Atlanta, GA.

2016 Milner GR, Boldsen JL, Ousley SD, Weise S, Getz SM, Tarp P. Improved Adult Age Estimation Using New Skeletal Traits and Transition Analysis. American Academy of Forensic Sciences, Las Vegas, NV.

INVENTIONS, PATENT APPLICATIONS, AND/OR LICENSES

Nothing to report for patents or licenses. The intent from the outset was to produce a skeletal age-estimation procedure for wide distribution to forensic practitioners without limitations upon its use. The computer program and supporting documentation is described elsewhere in this report.